THE DISTRIBUTION OF GENE FREQUENCIES IN POPULATIONS

BY SEWALL WRIGHT

DEPARTMENT OF ZOÖLOGY, UNIVERSITY OF CHICAGO

Read before the Academy, April 26, 1937

The effects of the various evolutionary factors—mutation, cross breeding, selection and inbreeding—can be reduced to common terms by considering the rates of change which they tend to bring about in the relative frequencies of alleles.[1] In the absence of such factors, there is constancy of gene frequencies from the symmetry of the Mendelian mechanism.

The frequency (q) of a given gene changes at the rate $\Delta q = -uq$ per generation under recurrent mutation of the gene to alleles at the rate u. Mutation in the opposite direction at the average rate v per generation changes the gene frequency at the rate, $\Delta q = v(1 - q)$.

If a certain gene has the frequency q in a local population but q_t in the species as a whole, exchange of the proportion m of the local population with an equal number of random individuals from the whole species leads to change of gene frequencies in the former at the rate $\Delta q = -m(q - q_t)$. Cross breeding is, however, most likely to be with neighboring populations which differ but little in value of q. In this case the coefficient m is only a small fraction of the actual amount of exchange. There may be other complications such as selective immigration or emigration, but the above simple form will suffice here to illustrate cross breeding or migration pressure.

The simplest kind of selection is that in which the heterozygote is exactly half way between the two homozygotes in the extent per individual to which it contributes to the next generation. The selective value of zygotes (relative to a certain standard) will be designated w and the mean value for a population, $\bar{w}$.

ZYGOTE	FREQUENCY	w	
AA	$(1 - q)^2$	1	$\bar{w} = 1 - 2sq$
AA'	$2q(1 - q)$	$1 - s$	$\frac{d\bar{w}}{dq} = -2s$
$A'A'$	q^2	$1 - 2s$	

$$\Delta q = \frac{(1 - s)q(1 - q) + (1 - 2s)q^2}{1 - 2sq} - q =$$

$$\frac{-sq(1 - q)}{1 - 2sq} = \frac{q(1 - q)}{2\bar{w}} \frac{d\bar{w}}{dq}. \qquad (1)$$

In the more general case in which $\bar{w}$ is not related linearly to q, the momentary selective advantage ($-2s$) of replacing A by A' is still given

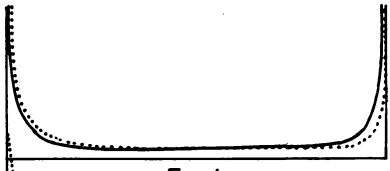

Fig. 1

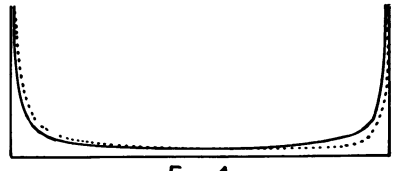

Fig. 4

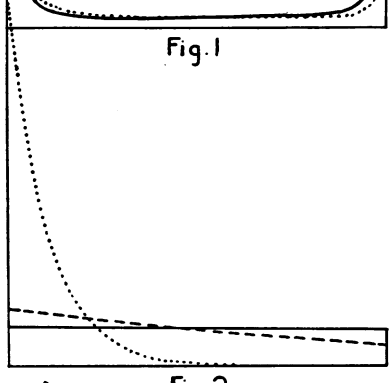

Fig. 2

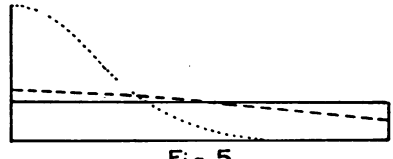

Fig. 5

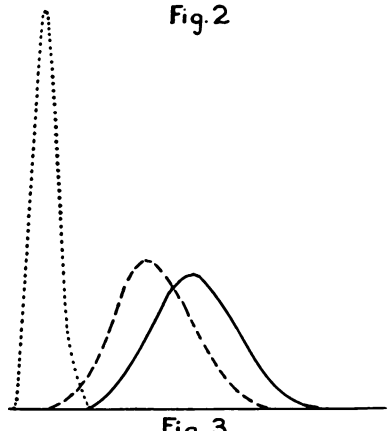

Fig. 3

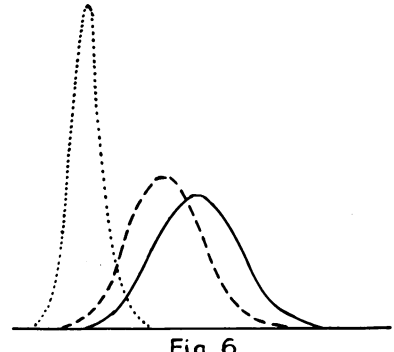

Fig. 6

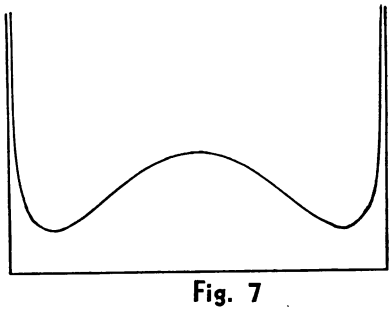

Fig. 7

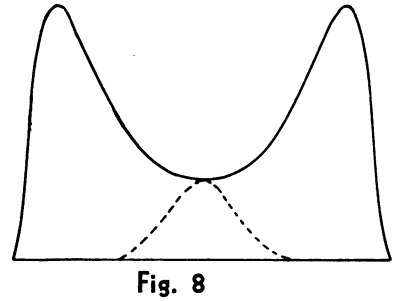

Fig. 8

(Captions for figures on opposite page.)

by $d\bar{w}/dq$ and the value of Δq is given by the same formula as above in terms of w, $d\bar{w}/dq$ and q.

ZYGOTE	FREQUENCY	w	
AA	$(1-q)^2$	1	$\bar{w} = 1 - 2s_1q(1-q) - s_2q^2$
AA'	$2q(1-q)$	$1-s_1$	$\frac{d\bar{w}}{dq} = -2s_1 + 2(2s_1 - s_2)q$
$A'A'$	q^2	$1-s_2$	

$$\Delta q = \frac{q(1-q)}{2\bar{w}}\frac{d\bar{w}}{dq}. \qquad (2)$$

Still more generally,[2] selective values depend on the interactions of the entire system of genes. It is the harmonious development of all characteristics that determines the success of an organism, not the absolute grades

CAPTIONS FOR FIGURES ON OPPOSITE PAGE

Figures 1 to 3. Some of the forms taken by the distribution of gene frequencies in the case of no dominance. $(\varphi(q) = Ce^{4Nsq}q^{4Nv-1}(1-q)^{4Nu-1})$. Mutation rates are assumed constant and equal $(u = v)$. Effective size of population is $N = \frac{1}{40v}, \frac{10}{40v}$ and $\frac{100}{40v}$ in figures 1, 2 and 3, respectively. In each case the solid line represents the least selection $(s = -v/10)$, the broken line selection 10 times as severe (not represented in figure 1 since practically indistinguishable from the preceding) and the dotted line represents selection 100 times as severe.

Figures 4 to 6. Some of the forms taken by the distribution of frequencies of a completely recessive deleterious gene. $Ce^{2Ntq^2}q^{2Nu-1}(1-q)^{4Nu-1}$. Mutation rates are assumed constant and equal $(u = v)$. Effective size of population is $N = \frac{1}{40v}, \frac{10}{40v}$ and $\frac{100}{40v}$ in figures 4, 5 and 6, respectively. In each case the solid line represents the least selection $(t = -v/5)$, the broken line selection 10 times as severe (not represented in figure 4 since practically indistinguishable from the preceding) and the dotted line represents selection 100 times as severe.

Figure 7. One of the forms taken by the distribution of gene frequencies in the case in which there is no adaptive difference between the two homozygous types but the heterozygote is selected over both. $Ce^{4Nsq(1-q)}q^{4Nu-1}(1-q)^{4Nv-1}$. In the case represented, $u = v$, $N = \frac{1}{40v}$, $s = 100v$.

Figure 8. The frequencies along two diagonals of the joint distribution for two series of alleles with equal and additive effects on a character on which adverse selection acts according to the square of the deviation from the mean. The solid line shows the frequencies in populations along the line connecting the two favorable types $A_1A_1a_2a_2$ and $a_1a_1A_2A_2$. The broken line refers to the line connecting the extreme types $a_1a_1a_2a_2$ and $A_1A_1A_2A_2$. $\varphi(q) = C[1 - 2s[q_1(1-q_1) + q_2(1-q_2) + 2(q_1+q_2-1)^2]^{2N}q_1^{4Nv_1-1}(1-q_1)^{4Nu_1-1}q_2^{4Nv_2-1}(1-q_2)^{4Nu_2-1}$. In the case shown, $u_1 = v_1 = u_2 = v_2$; $N = 1/2v_1$, $s = 5v_1$. Along the favorable diagonal $q_1 = (1-q_2)$ the distribution is approximately $Ce^{-20q_1(1-q_1)}q_1^2(1-q_1)^2$. Along the unfavorable diagonal $(q_1 = q_2)$ it is approximately $Ce^{-20(1-2q_1(1-q_1))}q_1^2(1-q_1)^2$.

of separate elementary characters, and still less its composition with respect to a single series of alleles. A gene that is favorable in one combination may be deleterious in another. However, if values of w are assigned to each possible combination, the rate of change of the frequency of a particular gene under selection (with specified values of all other gene frequencies) is given by the formula

$$\Delta q_i = \frac{q_i(1 - q_i)}{2\bar{w}} \frac{\partial \bar{w}}{\partial q_i}. \tag{3}$$

If the selection coefficients are small, $\bar{w}$ is close to 1, and the form $^1/_2 q_i(1 - q_i)\partial\bar{w}/\partial q_i$ is sufficiently accurate and is sometimes more convenient.

Two or more of the above factors are usually acting simultaneously. If the rates of change per generation are small, the net rate of change is given sufficiently accurately by the sum.

$$\Delta q = -uq + v(1 - q) - m(q - q_l) + \frac{q(1 - q)}{2} \frac{\partial}{\partial q} \log \bar{w}. \tag{4}$$

Gene frequency is in equilibrium (stable or otherwise) at any point at which $\Delta q = 0$. Opposing mutation pressures, for example, tend to maintain a stable equilibrium at the point $\hat{q} = v/(u + v)$. Mutation opposed by sufficiently strong genic selection [$\Delta q = v(1 - q) - sq(1 - q)$] gives stable equilibrium at the point $\hat{q} = v/s$. Recessive mutation opposed by sufficiently strong selection [$\Delta q = v(1 - q) - sq^2(1 - q)$] gives stable equilibrium at the point $\hat{q} = \sqrt{v/s}$. If there is mutation in both directions and sufficiently strong selection against the heterozygote [$\Delta q = -uq + v(1 - q) - s(1 - 2q)q(1 - q)$] there are two points of stable equilibrium, and one of unstable equilibrium.

Migration pressure, if non-selective, may be written in the same form as mutation pressure, $\Delta q = -m(1 - q_l)q + mq_l(1 - q)$. The theory for migration effects can thus be obtained at any time from that for mutation merely by substituting $m(1 - q_l)$ for u and mq_l for v.

If the population is not indefinitely large, random changes occur in gene frequencies merely as a result of the accidents of sampling among the gametes. Letting N be the effective size of the breeding population, the sample of $2N$ gametes, necessary to replace it, will be distributed according to the expansion of $[(1 - q)A + qA']^{2N}$. The resulting standard deviation of q is $\sqrt{q(1 - q)/2N}$.

The changes in gene frequency due to accidents of sampling are, of course, not correlated in successive generations. Nevertheless, the variance of the probability array for q increases approximately with the number of generations until damped by the approach of q to 0 or 1.

The pressure toward a stable equilibrium point due to mutation, cross breeding and selection, and the divergent tendency due to inbreeding should between them determine a certain *distribution* of values of q which is in equilibrium. The central problem in the genetics of populations is that of finding this distribution under various conditions.

The first attempt at a solution was made by Fisher[3] who used a transformation of scale, $\theta = \cos^{-1}(1 - 2q)$, designed to give a uniform sampling variance, for all values of q. He attempted to express the conditions in a differential equation but reached erroneous conclusions. My first note on the subject was in 1929,[4] the detailed account appearing in 1931.[1] Fisher (1930)[5,6] after inspection of the latter paper in manuscript was able to correct his method so as to yield results in agreement in a number of special cases.

The method followed in the 1931 paper, referred to above, may be summarized as follows. A class of genes with frequency array $[(1 - q)A + qA']$ is distributed in the following generation according to the expansion $[(1 - q - \Delta q)A + (q + \Delta q)A']^{2N}$. The contribution to the class of genes with frequency array $[(1 - q_c)A + q_cA']$ is given by the term in the expansion relating to $2Nq_cA''$s, multiplied by the frequency (f) of the contributing class. The sum of such contributions from all classes of genes should restore the same frequency as in the preceding generation if the form of distribution is in equilibrium. The distinction between gene frequency (q) and frequency (f) of a gene frequency must be kept in mind.

$$f_c = \frac{\lfloor 2N}{\lfloor 2Nq_c \lfloor 2N(1 - q_c)} \sum_{q=0}^{1} [(q + \Delta q)^{2Nq_c}(1 - q - \Delta q)^{2N(1 - q_c)}f]. \quad (5)$$

Replacing summation by integration and letting $f = \varphi(q)/2N$ or $\varphi(q)dq$ according to position in the equation, the equation to be solved for $\varphi(q)$ is as follows

$$\varphi(q_c) = \frac{\Gamma(2N)}{q_c(1 - q_c)\Gamma(2Nq_c)\Gamma(2N(1 - q_c))} \int_0^1 (q + \Delta q)^{2Nq_c}(1 - q - \Delta q)^{2N(1 - q_c)}\varphi(q)dq. \quad (6)$$

If Δq is negligibly small—except for enough mutation from the homallelic classes ($q = 0$ or 1) to make equilibrium possible—it may easily be seen that the solution of this equation is

$$\varphi(q) = \frac{C}{q} + \frac{D}{1 - q}. \quad (7)$$

Only the symmetrical case, however, is in complete equilibrium with the homallelic classes, although the rate of change in other cases is extremely slow.

$$\varphi(q) = \frac{(l)}{2(\log 2N + .577)q(1 - q)} \tag{8}$$

when (l) is the proportion of heterallelic loci. The proportions in the homallelic loci are

$$f(0) = \frac{1}{4Nv} f(1/2N),\ f(1) = \frac{1}{4Nu} f\,(1 - 1/2N). \tag{9}$$

$$f(0) + l + f(1) = 1$$

If there is no mutation, there is equilibrium of form among the heterallelic classes when all are approximately equally frequent but falling off at the rate $1/2N$ per generation.

$$\varphi(q) = (l_0 e^{-T/2N}) \tag{10}$$

where l_0 is the initial proportion in heterallelic loci and T is the number of subsequent generations. The two preceding results were confirmed by Fisher[5, 6] by his method.

If mutation is recurring at appreciable rates such that $\Delta q = -uq + v(1 - q)$, the distribution was shown to take the form

$$\varphi(q) = \frac{\Gamma(4Nu + 4Nv)}{\Gamma(4Nu)\Gamma(4Nv)} q^{4Nv-1}(1 - q)^{4Nu-1}. \tag{11}$$

With irreversible mutation, $\Delta q = v(1 - q)$, there is equilibrium of form, with falling off of all class frequencies at the rate v per generation

$$\varphi(q) = (l_0 e^{-vT}) 4Nv q^{4Nv-1}. \tag{12}$$

With genic selection but very small mutation rates, $\Delta q = sq(1 - q)$, equilibrium of form was shown for the distribution

$$\varphi(q) = \frac{Ce^{4Nsq} + D}{q(1 - q)}. \tag{13}$$

The case

$$\frac{Ce^{4Nsq}}{q(1 - q)} \tag{14}$$

is that which is in complete equilibrium with the homallelic classes. Another case is the solution

$$\varphi(q) = \frac{C(e^{4Nsq} - 1)}{q(1 - q)} \tag{15}$$

given by Fisher[5,6] for irreversible mutation.

It was shown finally that the formula for mutation and genic selection combined, $\Delta q = -uq + v(1 - q) + sq(1 - q)$, could be written sufficiently accurately

$$\varphi(q) = Ce^{4Nsq}q^{4Nv-1}(1 - q)^{4Nu-1}. \qquad (16)$$

The effects of cross breeding could be introduced in place of (or supplementary to) mutation by the substitution previously referred to.

Haldane[7] has criticized the conclusions derived from these formulae

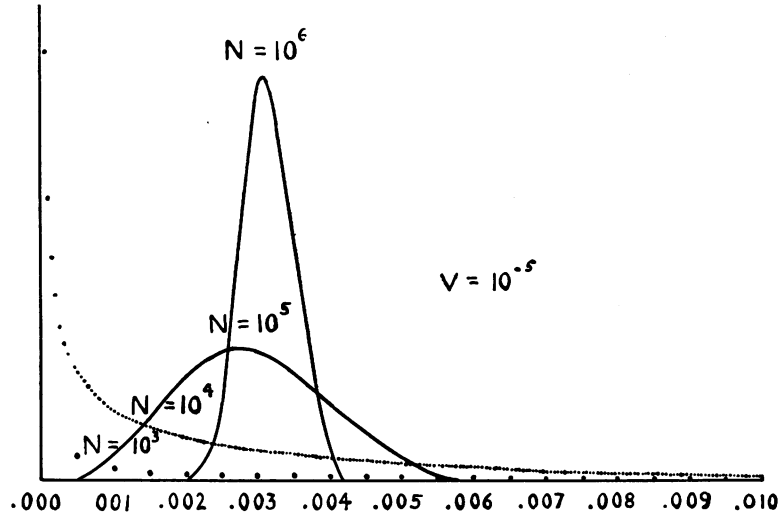

Figure 9. Some of the forms taken by the distribution of frequencies of a recessive lethal gene under different sizes of population. $C(1 - q^2)^{2N}q^{4Nv-1}(1 - q)^{-1}$. The mutation rate ($v$) is taken as 10^{-5}, giving a mean gene frequency of $\sqrt{v} = .0032$ in very large populations. This is approximately realized with $N = 10^6$. With $N = 10^5$, the gene is always present, but $\bar{q}$ is slightly reduced (.0030). In populations with effective size, $N = 10^4$, the gene is absent in about 15%, at any given moment and $\bar{q} = .0020$. With $N = 10^3$, the gene is absent in 87%, and $\bar{q} = .0008$. The case of $N = 10^2$ is not shown as the gene is absent in about 99% and $\bar{q}$ is only .00026. With $N = 10$, the gene is absent in 99.9%, and $\bar{q} = .00008$. In selfed lines ($N = 1$) the gene is absent in 99.996% giving $\bar{q} = .00002$ ($= 2v$).

on the ground that only one type of selection is considered. It is undoubtedly important to extend these results to more general formulations of the action of selection. This has now been done by the same method as above for selection pressure of the form $\Delta q = (s + tq)q(1 - q)$ which applies to any degree of dominance, assuming s and t to be small. I have not been able to reduce the general solution to a simpler form than an infinite series, but in the case of complete equilibrium with the fixed classes there is reduction to the following

$$\varphi(q) = \frac{Ce^{4Nsq + 2Ntq^2}}{q(1 - q)}. \qquad (17)$$

Still more general results can be obtained by another approach (presented as an alternative in the preceding paper in the case of mutation pressure). A frequency distribution, ranging from 0 to 1, in which mean and standard deviation do not change under evolutionary pressure and sampling errors, and whose mode (or modes) can be shown to be correct in the limiting case of large population size, must be a good approximation to the desired type. Let $\varphi(q)$ be any distribution of gene frequencies.

Then

$$\bar{q} = \int_0^1 q\varphi(q)dq \text{ (assuming that } \int_0^1 \varphi(q)dq = 1). \qquad (18)$$

Accidents of sampling by themselves can have no effect in changing the mean, the mean of the distribution $[(1 - q)A + qA']^{2N}$ being $2Nq$. The effect of any evolutionary pressure in changing the mean is given by the expression

$$\int_0^1 \Delta q\varphi(q)dq. \qquad (19)$$

If Δq is linear in q, and therefore of the form $\Delta q = -K(q - \hat{q})$ where $\hat{q}$ is the equilibrium point, the change of mean is

$$\int_0^1 \Delta q\varphi(q)dq = -K\int_0^1 (q - \hat{q})\varphi(q)dq = -K(\bar{q} - \hat{q}). \qquad (20)$$

Thus if the mean of $\varphi(q)$ has reached the equilibrium point ($\bar{q} = \hat{q}$) there is no further change. Mutation pressure, $\Delta q = -uq + v(1 - q) = -(u + v)(q - \hat{q})$, comes under this head. The mean of the distribution of gene frequencies is always $\hat{q} = v/(u + v)$ at equilibrium, irrespective of size of population. Cross breeding also comes under this head, $\Delta q = -m(q - q_t)$. Selection pressure on the other hand, with Δq at least a quadratic function of q should give different values of $\bar{q}$ with change in size of population.

The variance of gene frequencies is also easily found in the case of linear evolutionary pressures. The pressure from both directions toward the equilibrium point must be balanced by the scattering effect of accidents of sampling if there is equilibrium in the form of distribution.

$$\sigma_q^2 = \int_0^1 (q - \bar{q})^2\varphi(q)dq = \int_0^1 (q + \Delta q - \bar{q})^2\varphi(q)dq + \int_0^1 \frac{q(1 - q)}{2N}\varphi(q)dq. \qquad (21)$$

Inserting the value $\Delta q = -K(q - \bar{q})$ this reduces to

$$\sigma_q^2 = \frac{\bar{q}(1 - \bar{q})}{1 + 4NK - 2NK^2}. \qquad (22)$$

The term in the denominator involving K^2, tracing to the term in (21) involving $(\Delta q)^2$ is negligible for small values of Δq.

Thus if $\Delta q = -K(q - \bar{q})$, the distribution of gene frequency has the mean, $\bar{q}$; variance, $\sigma_q^2 = \frac{\bar{q}(1 - \bar{q})}{1 + 4NK}$; a range limited at $q = 0$ and $q = 1$ and for large values of N condenses about a single mode at the equilibrium point, $q = \bar{q}$. In Pearson's system of frequency distributions, unimodal curves of limited range come under type I. Assuming a range limited at 0 and 1 and a unit area, the formula for type I is $\varphi(q) = \frac{\Gamma(x + y)}{\Gamma(x)\Gamma(y)} q^{x-1}(1 - q)^{y-1}$. Substituting this value of $\varphi(q)$ in the formulae for $\bar{q}$ and σ_q^2 gives $x = 4NK\bar{q}$, $y = 4NK(1 - \bar{q})$. In the case of mutation, $K = u + v$, $\bar{q} = v/(u + v)$.

$$\varphi(q) = \frac{\Gamma(4Nu + 4Nv)}{\Gamma(4Nu)\Gamma(4Nv)} q^{4Nv-1}(1 - q)^{4Nu-1}. \qquad (23)$$

In the case of migration, $K = m$, $\bar{q} = q_t$

$$\varphi(q) = \frac{\Gamma(4Nm)}{\Gamma(4Nmq_t)\Gamma[4Nm(1 - q_t)]} q^{4Nmq_t-1}(1 - q)^{4Nm(1-q_t)-1}. \qquad (24)$$

These are identical with the formulae derived by the preceding method (cf. 11).

When selection as well as mutation is at work we may write

$$\Delta q = Zq(1 - q) - uq + v(1 - q) \text{ where } Z = {}^1/_2\, d \log \bar{w}/dq. \qquad (25)$$

A suggestion for the distribution formula can be obtained from the special case in which N is so large that there is very little spread from the equilibrium point (or points) at $\Delta q = 0$. The modes should approach the equilibrium points as N is increased. We will make the provisional assumption that the selection coefficients appear in a separate factor from those involving the mutation coefficients and that the desired formula is therefore of the type $\varphi(q) = \psi q^{4Nv-1}(1 - q)^{4Nu-1}$ where ψ is the required function of the selection coefficient. Putting $d \log \varphi(q)/dq = 0$ as a condition for any mode:

$$\frac{d}{dq} \log \varphi(q) = \frac{d}{dq} \log \psi + \frac{4Nv - 1}{q} - \frac{4Nu - 1}{1 - q} = 0. \qquad (26)$$

Thus

$$\frac{q(1-q)}{4N}\frac{d}{dq}\log\psi + v(1-q) - uq + \frac{2q-1}{4N} = 0. \tag{27}$$

But at the equilibrium point,

$$Zq(1-q) + v(1-q) - uq = 0. \tag{28}$$

Ignoring the term $(2q-1)/4N$ which tends to disappear as N becomes large

$$Z = \frac{1}{4N}\frac{d}{dq}\log\psi. \tag{29}$$

$$\psi = Ce^{4N\int Zdq}. \tag{30}$$

The formula suggested on this basis is thus as follows:

$$\varphi(q) = Ce^{4N\int Zdq}q^{4Nv-1}(1-q)^{4Nu-1}. \tag{31}$$

This can be written as follows by evaluating $\int Zdq$

$$\varphi(q) = C\bar{w}^{2N}q^{4Nv-1}(1-q)^{4Nu-1}. \tag{32}$$

We will return to this form later; for the moment it will be convenient to use the following alternative form, easily derived from (25) and (31).

$$\varphi(q) = \frac{Ce^{4N\int\Delta qdq/q(1-q)}}{q(1-q)}. \tag{33}$$

So far this is merely a suggestion derived from a limiting case. It may be tested, however, for equilibrium in the general case. Testing first for shifting of the mean under evolutionary pressure (see 19)

$$\int_0^1 \Delta q\varphi(q)dq = \frac{C}{4N}\int_0^1 \frac{4N\Delta q}{q(1-q)}e^{\int 4N\Delta qdq/q(1-q)}dq =$$

$$\frac{C}{4N}\Big|_0^1 e^{4N\int Zdq}q^{4Nv}(1-q)^{4Nu}. \tag{34}$$

Thus there is no shifting of the mean if $4Nv > 0$, $4Nu > 0$.

The test (21) for balancing of the effects on variance of evolutionary pressure and accidents of sampling can be written as follows, omitting the negligible term in $(\Delta q)^2$.

$$2\int_0^1 q\Delta q\varphi(q)dq = -\frac{1}{2N}\int_0^1 q(1-q)\varphi(q)dq. \tag{35}$$

The left-hand member can be integrated by parts after substituting the suggested value of $\varphi(q)$ and shown to equal the right-hand member.

$$2C\int_0^1 \frac{q\,\Delta q}{q(1-q)}\, e^{4N\int \Delta q dq/q(1-q)}dq =$$

$$\frac{C}{2N}\int_0^1 qd(e^{4N\int \Delta q dq/q(1-q)}) = \frac{C}{2N}\Big|_0^1 e^{4N\int Z dq} q^{4Nv+1}(1-q)^{4Nu} -$$

$$\frac{C}{2N}\int_0^1 e^{4N\int \Delta q dq/q(1-q)}dq = -\frac{1}{2N}\int_0^1 q(1-q)\varphi(q)dq. \qquad (36)$$

The condition of no change in variance is thus met to the same degree of approximation as in the case of mutation. The special case of mutation and genic selection, $\Delta q = -uq + v(1-q) + sq(1-q)$ gives $\varphi(q) = Ce^{4Nsq}q^{4Nv-1}(1-q)^{4Nu-1}$, identical with that (16) derived by the previous more exhaustive method. This holds even in the limiting case (14) in which $\Delta q = sq(1-q)$ but as already noted equilibrium requires that there be some mutation even though the rates are so low that they may be ignored in Δq. The formula (17), obtained by the more exhaustive method in the case of more or less dominance, $\Delta q = (s + tq)q(1-q)$ is also in agreement with the present result.

The effects of certain differences in severity of selection and in effective size of population on the distribution of gene frequencies, assuming no dominance, are illustrated in figures 1 to 3. Figures 4 to 6 make similar comparisons for the case of a completely recessive unfavorable gene. These figures can also illustrate the joint effects of selection, cross breeding and inbreeding in local populations by replacing u and v by $\frac{1}{2}m$.

One of the forms taken by the distribution when there is equal selection against both homozygotes in favor of the heterozygote is illustrated in figure 7. With sufficiently smaller population size or sufficiently less intense selection, the distribution would become U-shaped. With a sufficiently larger population size it would become I-shaped about a mean frequency, $q = 0.5$. Increased severity of selection, without increase in size of population, would also pile up the frequencies about this point.

In a large population, mutation opposed by moderately severe selection tends to hold the deleterious gene at a low frequency, $\hat{q} = v/s$ if semi-dominant, $\hat{q} = (v/s)^{1/2}$ if completely recessive. In a sufficiently small sample from such a population, selection of the same degree of severity becomes ineffective and the mean frequency of the deleterious gene gradually rises to the equilibrium point due to opposing mutation pressures $\hat{q} = v/(u+v)$. If unfavorable mutation is much more frequent than the reverse ($v > u$) this may mean a shift to approximate fixation of the deleterious gene. The rate of approach to the new mean, after a reduction in size of population is, however, extremely slow, being dependent on mutation pressure.

In the case of a deleterious *recessive* factor, the immediate effect of reduction in size of population is indeed the reverse of that indicated above, though the final effect, a rise in mean frequency of the deleterious gene toward the point $q = v/(u + v)$ occurs at length in this case as well as when dominance is lacking (provided that populations in which the deleterious gene is fixed can persist). The immediate decrease in frequency on reduction of size of population is illustrated in figure 9 in the extreme case of a recessive lethal in which case there can, of course, be no secondary rise in frequency. Taking the dominant as the type, $\bar{w} = 1 - q^2$, $\Delta q = v(1 - q) - q^2/(1 + q)$ giving the approximate equilibrium point $q = v^{1/2}$. The distribution is

$$\varphi(q) = C(1 - q^2)^{2N} q^{4Nv - 1}(1 - q)^{-1}. \qquad (37)$$

The mean, $\bar{q}$, can easily be expressed in Γ functions (on substituting x for q^2). It reduces approximately to $\bar{q} = \dfrac{\Gamma(2Nv + 1/2)}{\sqrt{2N}\,\Gamma(2Nv)}$. For values of $2Nv$ larger than 1, this is close to $v^{1/2}$ i.e., gene frequency varies about the equilibrium point. If on the other hand $2Nv$ is a small fraction, q is approximately $v(2\pi N)^{1/2}$ which means a great reduction in frequency of lethals in populations as the effective size of inbreeding units falls below $1/2v$.

So far we have derived formulae for distributions only where the selection coefficients are constant. But as already noted, it is really the system of gene frequencies that is more or less adaptive, not the isolated genes. No adequate picture of the evolutionary process can be made without taking factor interaction into account.

The momentary selection pressure in cases of factor interaction was given in (3), giving as the distribution of gene frequencies

$$\varphi(q_i) = C\bar{w}^{2N} q_i^{4Nv_i - 1}(1 - q_i)^{4Nu_i - 1}. \qquad (38)$$

Here $\bar{w}$ is the mean selective value with q_i variable, but a specified set of values for the other gene frequencies.

The joint frequency surface must be such that on assigning specified values to all of the q's but one, the distribution for that one is that given by (38). The joint distribution

$$\varphi(q_1, q_2, \ldots q_n) = C\bar{w}^{2N}\Pi_{i=1}^{n} q_i^{4Nv_i - 1}(1 - q_i)^{4Nu_i - 1} \qquad (39)$$

when $\bar{w}$ is the mean selective value in terms of all of the q's as variables, satisfies this condition and is thus the desired form.

As an example consider the case of a character for which the grade of development depends on the additive effects of multiple factors, lacking dominance, but for which the selective value falls off as the square of its deviation from an optimum. The selective value has been shown to be as follows:[2]

$$\bar{w} = 1 - K[2\Sigma\alpha_i^2 q_i(1 - q_i) + (M - O)^2]$$

where α_i is the effect of gene A_i with frequency q_i; $M(= 2\Sigma\alpha_i q_i)$ is the mean and O is the optimal grade.

The multidimensional surface, $\bar{w}$, has in general many peaks, separated by shallow saddles. The shallowest saddles are those between optimal combinations differing only in two pairs of factors, e.g., $A_1A_1a_2a_2\ldots\ldots$ and $a_1a_1A_2A_2\ldots\ldots$

The nature of the distribution along and across such a saddle is illustrated in figure 8 in the case of two pairs of factors with equal effects. With smaller N the distribution from $A_1A_1a_2a_2$ to $a_1a_1A_2A_2$ becomes U-shaped. With larger N or weaker selection it becomes I-shaped about $q = 0.5$. Stronger selection pushes the modes toward the favored homallelic types.

The evolutionary implications will not be discussed here in detail. For the most part the present results merely put the conclusions previously reached[1,8,9] on a more definite basis. These conclusions may be summarized briefly as follows.

In large freely interbreeding populations with no secular change in conditions of life for long periods of time, all gene frequencies approach equilibrium at a certain peak $\bar{w}$, not necessarily the highest peak. Under secular change in conditions the surface $\bar{w}$ itself changes and there is evolutionary change in the system of gene frequencies, following the changes in position of the controlling peak. Evolution here may be said to be guided by intragroup selection.

In sufficiently small completely isolated populations, the random divergencies of gene frequencies from their equilibrium values become important, tending to bring about approximate fixation of some random combination of genes which is not likely to be a peak combination. The result is a largely nonadaptive differentiation. In extreme cases there may be the deterioration which characteristically follows excessive inbreeding. Isolation may here be considered the dominating evolutionary factor.

In a large population subdivided into numerous small, partially isolated groups, the combination of directed and random divergencies in gene frequencies, associated with intergroup selection, gives a trial and error mechanism under which the system of gene frequencies may pass from lower to higher peak values of $\bar{w}$ and the species may evolve continuously even without secular changes in conditions (although this process, occurring in all species, itself tends to bring about such secular changes). The combination of partial isolation of subgroups with intergroup selection seems to provide the most favorable conditions for evolutionary advance.

Mutation is always a factor in providing material for evolution but may be said to dominate the course of evolution only in so far as mutants

appear which are fertile *inter se* but largely infertile with the parent type, i.e., when mutation is itself an isolating factor.

[1] Wright, S., *Genetics*, **16**, 97–159 (1931).
[2] Wright, S., *Jour. Genetics*, **30**, 243–256 (1934).
[3] Fisher, R. A., *Proc. Roy. Soc. Edinburgh*, **42**, 321–341 (1922).
[4] Wright, S., *Amer. Nat.*, **63**, 556–561 (1929).
[5] Fisher, R. A., *The Genetical Theory of Natural Selection*, Oxford, Clarendon Press, 272 pp. (1930).
[6] Fisher, R. A., *Proc. Roy. Soc. Edinburgh*, **50**, 205–220 (1930).
[7] Haldane, J. B. S., *The Causes of Evolution*, New York and London, Harper Brothers, 235 pp. (1932).
[8] Wright, S., *Proc. Sixth Internat. Congress of Genetics*, **1**, 356–366 (1932).
[9] Wright, S., *Jour. Genetics*, **30**, 257–266 (1935).

THE INFLUENCE OF WAVE-LENGTH ON GENETIC EFFECTS OF X-RAYS[1]

BY HUGO FRICKE AND M. DEMEREC

THE WALTER B. JAMES LABORATORY FOR BIOPHYSICS OF THE LONG ISLAND BIOLOGICAL ASSOCIATION AND THE DEPARTMENT OF GENETICS OF THE CARNEGIE INSTITUTION OF WASHINGTON, COLD SPRING HARBOR, N. Y.

Communicated May 13, 1937

A great deal of work has been done with genetic effects of x-ray radiation. It has been shown that this effect, as measured by lethals produced in mature sperm of *Drosophila melanogaster*, is proportional to the dosage in r-units. It has also been found that the dosage-genetic effect relationship is not affected by the wave-length within a range of 0.02 to 2.0 Å.

It is a problem of theoretical importance to determine whether or not the mechanism which produces genetic effects is dependent upon wave-length, especially within the range of soft rays. In studying this range, however, technical difficulties due to high absorption are encountered.

The purpose of this work was to obtain data on the genetic effect of soft rays in experiments where both physical and biological sides were well controlled. A particular effort was made to control the absorption. While these experiments were in progress, results of similar experiments conducted by Timofeeff-Ressovsky and Zimmer[2] were published. Results of both experiments agree, although there is a disagreement in their interpretation.

Experimental Procedure. Physical Part.—The x-rays were obtained from a tungsten tube with a bulb of lithium glass having a 0.02 mm. window for the exit of the soft rays. The conditions of irradiation are given in table 1. The quality of the rays was determined by measuring their ab-